



**University of  
Zurich**<sup>UZH</sup>

**Zurich Open Repository and  
Archive**

University of Zurich  
University Library  
Strickhofstrasse 39  
CH-8057 Zurich  
[www.zora.uzh.ch](http://www.zora.uzh.ch)

---

Year: 2014

---

## **Breath analysis in real time by mass spectrometry in chronic obstructive pulmonary disease**

Martinez-Lozano Sinues, Pablo ; Meier, Lukas ; Berchtold, Christian ; Ivanov, Mark ; Sievi, Noriane ; Camen, Giovanni ; Kohler, Malcolm ; Zenobi, Renato

**Abstract:** BACKGROUND: It has been suggested that exhaled breath contains relevant information on health status. OBJECTIVES: We hypothesized that a novel mass spectrometry (MS) technique to analyze breath in real time could be useful to differentiate breathprints from chronic obstructive pulmonary disease (COPD) patients and controls (smokers and nonsmokers). METHODS: We studied 61 participants including 25 COPD patients [Global Initiative for Obstructive Lung Disease (GOLD) stages I-IV], 25 nonsmoking controls and 11 smoking controls. We analyzed their breath by MS in real time. Raw mass spectra were then processed and statistically analyzed. RESULTS: A panel of discriminating mass-spectral features was identified for COPD (all stages; n = 25) versus healthy nonsmokers (n = 25), COPD (all stages; n = 25) versus healthy smokers (n = 11) and mild COPD (GOLD stages I/II; n = 13) versus severe COPD (GOLD stages III/IV; n = 12). A blind classification (i.e. leave-one-out cross validation) resulted in 96% sensitivity and 72.7% specificity (COPD vs. smoking controls), 88% sensitivity and 92% specificity (COPD vs. nonsmoking controls) and 92.3% sensitivity and 83.3% specificity (GOLD I/II vs. GOLD III/IV). Acetone and indole were identified as two of the discriminating exhaled molecules. CONCLUSIONS: We conclude that real-time MS may be a useful technique to analyze and characterize the metabolome of exhaled breath. The acquisition of breathprints in a rapid manner may be valuable to support COPD diagnosis and to gain insight into the disease.

DOI: <https://doi.org/10.1159/000357785>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-102002>

Journal Article

Published Version

Originally published at:

Martinez-Lozano Sinues, Pablo; Meier, Lukas; Berchtold, Christian; Ivanov, Mark; Sievi, Noriane; Camen, Giovanni; Kohler, Malcolm; Zenobi, Renato (2014). Breath analysis in real time by mass spectrometry in chronic obstructive pulmonary disease. *Respiration*, 87(4):301-310.

DOI: <https://doi.org/10.1159/000357785>

# Breath Analysis in Real Time by Mass Spectrometry in Chronic Obstructive Pulmonary Disease

Pablo Martinez-Lozano Sinues<sup>a</sup> Lukas Meier<sup>a</sup> Christian Berchtold<sup>a</sup>  
Mark Ivanov<sup>d</sup> Noriane Sievi<sup>b</sup> Giovanni Camen<sup>b</sup> Malcolm Kohler<sup>b, c</sup>  
Renato Zenobi<sup>a</sup>

<sup>a</sup>Department of Chemistry and Applied Biosciences, ETH Zurich, <sup>b</sup>Pulmonary Division, University Hospital Zurich, and <sup>c</sup>Zurich Center for Integrative Human Physiology, University of Zurich, Zurich, Switzerland; <sup>d</sup>Institute for Energy Problems of Chemical Physics, Russian Academy of Science, Moscow, Russian Federation

## Key Words

Breath test · Chronic obstructive pulmonary disease · Exhaled biomarkers

## Abstract

**Background:** It has been suggested that exhaled breath contains relevant information on health status. **Objectives:** We hypothesized that a novel mass spectrometry (MS) technique to analyze breath in real time could be useful to differentiate breathprints from chronic obstructive pulmonary disease (COPD) patients and controls (smokers and non-smokers). **Methods:** We studied 61 participants including 25 COPD patients [Global Initiative for Obstructive Lung Disease (GOLD) stages I–IV], 25 nonsmoking controls and 11 smoking controls. We analyzed their breath by MS in real time. Raw mass spectra were then processed and statistically analyzed. **Results:** A panel of discriminating mass-spectral features was identified for COPD (all stages;  $n = 25$ ) versus healthy nonsmokers ( $n = 25$ ), COPD (all stages;  $n = 25$ ) versus healthy smokers ( $n = 11$ ) and mild COPD (GOLD stages I/II;  $n = 13$ ) versus severe COPD (GOLD stages III/IV;  $n = 12$ ). A blind classification (i.e. leave-one-out cross validation) resulted in 96% sensitivity and 72.7% specificity (COPD vs.

smoking controls), 88% sensitivity and 92% specificity (COPD vs. nonsmoking controls) and 92.3% sensitivity and 83.3% specificity (GOLD I/II vs. GOLD III/IV). Acetone and indole were identified as two of the discriminating exhaled molecules. **Conclusions:** We conclude that real-time MS may be a useful technique to analyze and characterize the metabolome of exhaled breath. The acquisition of breathprints in a rapid manner may be valuable to support COPD diagnosis and to gain insight into the disease.

© 2014 S. Karger AG, Basel

## Introduction

Chronic obstructive pulmonary disease (COPD) is a complex illness, which is sometimes difficult to properly diagnose and for which many fundamental questions remain unanswered [1]. There is a clear need to increase our knowledge of the COPD phenotype for the ultimate purpose of improving diagnostics [2].

Taking samples from the lung itself (e.g. induced sputum) to perform analytical measurements can be difficult, and for this reason the analysis of exhaled breath is eliciting considerable interest as it offers the unique advantage

of probing the lung itself, but in a noninvasive fashion [3, 4]. As a result, in recent years, a number of techniques have been developed to analyze exhaled breath for investigating COPD among other lung diseases. These include the analysis of exhaled-breath condensate [5–12], exhaled particles [13] and volatile compounds [14–20]. Even trained dogs seem to be able to recognize the peculiar scent of COPD in exhaled breath [21].

A new mass spectrometry (MS) technique has been developed which aims to detect trace compounds in breath or any other matrix in real time [22–28]. This technique has been termed secondary electrospray ionization MS (SESI-MS, also dubbed ‘extractive electrospray ionization’) [29], and in a clinical context, it has shown promise for the diagnosis of bacterial infections based on their emitted volatile ‘breathprints’ [30–34]. This study seeks to expand SESI-MS applications in order to develop noninvasive techniques to characterize COPD through exhaled breath. In particular, the aim of this exploratory study was to assess, for the first time, whether or not SESI-MS holds promise to be able to identify subjects with COPD. Given that this technique has never been used to study any lung disease, the type of response of SESI-MS to be expected in such cases was completely unknown. For this reason, we deliberately chose to include a young and healthy control group. The aim was to establish the lower and upper bounds to be expected in future measurements with age-matched controls, similar to previous pilot studies on breath analysis [35]. We nevertheless attempted to characterize potential differences between COPD patients with the mild or moderate stages [Global Initiative for Obstructive Lung Disease (GOLD) I/II] and the severe or very severe stages (GOLD III/IV) of the disease. We therefore analyzed the breath of this heterogeneous group and applied statistical methods to reveal whether or not differences in breath patterns would be observable in such cases.

## Material and Methods

### Subjects

A group of 61 subjects participated in this study: 25 healthy nonsmokers, 11 healthy asymptomatic smokers and 25 patients affected by COPD.

Inclusion criteria for subjects with COPD were: males and females aged between 40 and 75 years, objectively diagnosed with COPD according to the GOLD guidelines [36]. Exclusion criteria were: having (1) a mental or physical disability precluding informed consent or compliance with the protocol, and (2) an acute exacerbation of COPD within the last 6 weeks.

Inclusion criteria for nonsmoking controls were: males and females aged between 18 and 75 years with no clinical signs of lung

**Table 1.** Characteristics of study subjects

	GOLD III/IV (n = 12)	GOLD I/II (n = 13)	Healthy smokers (n = 11)	Healthy nonsmokers (n = 25)
Age, years	62±5	63±7	34±10	27±9
Male/female	3/9	9/4	9/2	18/7
Smoking status				
Non/ex/current	1/10/1	1/8/4	0/0/11	24/1/0
Pack-years	37±23	40±32	10±13	0
BMI	23±4	31±7	28±4	23±3
FEV <sub>1</sub> % predicted	27±6	70±13	111±10	101±9
FEV <sub>1</sub>	0.6±0.1	2.1±0.6	4.5±0.6	4.1±0.8
FVC % predicted	74±15	92±15	111±12	105±10
FVC	2.1±0.5	3.4±1	5.4±0.9	5±1.1
FEV <sub>1</sub> /FVC	31±7	63±6	83±4	82±7
GOLD stage				
I/II/III/IV	2/10	2/11	0	0
Values are mean ± SD.				

disease. Exclusion criteria were: (1) having a mental or physical disability precluding informed consent or compliance of the protocol; (2) being pregnant, and (3) being a current smoker.

Inclusion criteria for smoking controls were: (1) males and females aged between 18 and 75 years with no clinical signs of lung disease, and (2) being a current smoker. Exclusion criteria were: (1) having a mental or physical disability precluding informed consent or compliance with the protocol, and (2) being pregnant.

The COPD stage (i.e. severity of airflow limitation) was assessed based on post-bronchodilator forced expiratory volume in 1 s (FEV<sub>1</sub>). As a result, 13 COPD patients were diagnosed as mild or moderate (GOLD stages I/II) and 12 were diagnosed as severe or very severe (GOLD stages III/IV). Two out of the 25 COPD patients had never smoked, while 18 were ex-smokers and 5 were still smoking.

The participants were asked not to smoke, eat, drink (except water) or brush their teeth (including the use of mouthwash or sprays) for at least 1 h before the measurements were performed. Subjects were randomly allocated to measurement slots in order to prevent any bias due to instrument drift.

The study was carried out in accordance with the Declaration of Helsinki of the World Medical Association. The local ethics committee approved the study (COPD patients EK-NR: 2011-0106 and healthy controls: EK 2012-N-49) and all subjects gave their written informed consent to participate. Measurements took place between March 3 and June 5, 2012.

### Mass Spectrometry Measurements

The experimental set-up was previously presented in detail [22, 24, 26, 28, 30–34, 37]. Briefly, the entrance of a commercial quadrupole time-of-flight mass spectrometer (Ultima, Waters, Milford, Mass., USA) was slightly modified to allow for the admission of breath samples through a heated Teflon tube. Exhaled breath was mixed with a nanoelectrospray plume (water and 0.2% formic acid), whereby some compounds present in the breath were read-

**Table 2.** Study subjects' medication

Medication	GOLD III/IV (n = 12)	GOLD I/II (n = 13)	Healthy smokers (n = 11)	Healthy non-smokers (n = 25)
Short-acting beta-2-sympathomimetic	8	1	0	0
Long-acting beta-2-sympathomimetic	11	5	0	0
Short-acting anticholinergic	1	0	0	0
Long-acting anticholinergic	10	4	0	0
Inhalational steroid	11	4	0	0
Antihypertensive medication	5	9	0	0
Antidiabetic medication	1	1	0	0
Statin	4	4	0	0
Long-term oxygen	1	0	0	0

ily detected in real time. The mass spectrometer was operated in the positive-ion mode. The subjects were asked to provide a deep exhalation through a disposable mouth piece, while keeping the pressure through the sampling tube at 20 mbar (monitored by a digital manometer), thereby ensuring that each subject breathed at the same flow rate (approx. 1.8 liters/min). This process was repeated 6 times per subject, with the 6 replicate measurements taking typically  $\leq 10$  min. Online supplementary figure S1 ([www.karger.com/doi/10.1159/000357785](http://www.karger.com/doi/10.1159/000357785)) shows a picture of a SESI-MS breath analysis set-up.

#### Statistical Analysis

The mass spectra, comprising the last few seconds of each exhalation, which reflects the exhaled composition of the lower respiratory tract, were averaged via Masslynx (i.e. Water's software) and exported as text files. These raw mass spectra were then processed using the commercial software MATLAB (R2012a, Mathworks Inc., Natick, Mass., USA). The original mass spectra were interpolated to 10,000 mass-to-charge ( $m/z$ ) values (56–400 Da; binned to  $\Delta m/z = 0.0187$  Da). The spectra were normalized by standardizing the area under the curve to the median. The local maxima of each peak were identified, resulting in 1,260 features with a threshold intensity of 10 counts. The 6 replicate measurements of each subject were averaged and, finally, a  $61 \times 1,260$  matrix was assembled. The 61 participants were allocated to 1 of the following groups: healthy nonsmokers, healthy smokers, mild COPD (GOLD I/II) and severe COPD (GOLD III/IV).

The analysis was conducted by pairs of groups separately: COPD (all stages;  $n = 25$ ) versus healthy nonsmokers ( $n = 25$ ), COPD (all stages;  $n = 25$ ) versus healthy smokers ( $n = 11$ ) and mild COPD (GOLD I/II;  $n = 13$ ) versus severe COPD (GOLD III/IV;  $n = 12$ ).

We first sought to reduce the dimensionality of the matrices by extracting the most significant features to discern these 4 groups. In this way, the noisiest peaks stemming from background air, chemical noise, etc. were largely excluded from the subsequent analysis. To do so, a Mann-Whitney U test was performed, looping each of the 1,260  $m/z$  values [38]. To prevent the inclusion of the noisiest peaks in further analysis, we did not consider the signal if the median for the 2 groups compared was  $< 50$  counts. Upon performance of the Mann-Whitney U test, only the signals with  $p$  values  $< 0.05$  (chosen significance level) were retained (4 for mild vs.

severe, 37 for COPD vs. nonsmoking controls and 69 for COPD vs. smoking controls).

Thus, the original data was reduced to matrices of  $25 \times 4$  (GOLD I/II vs. GOLD III/IV),  $36 \times 69$  (COPD vs. smoking controls) and  $50 \times 37$  (COPD vs. nonsmoking controls). To enable visualization of these data, the matrices were autoscaled [39] and subjected to principal component analysis [40]. Finally, the prediction power of these subsets of signals was assessed. A  $k$ -nearest neighbor ( $k$ -NN; Euclidean distance) classification algorithm was used to predict the class of a given breath spectrum in a leave-one-out cross validation exercise (i.e. each of the samples was left out of the model once and classified).

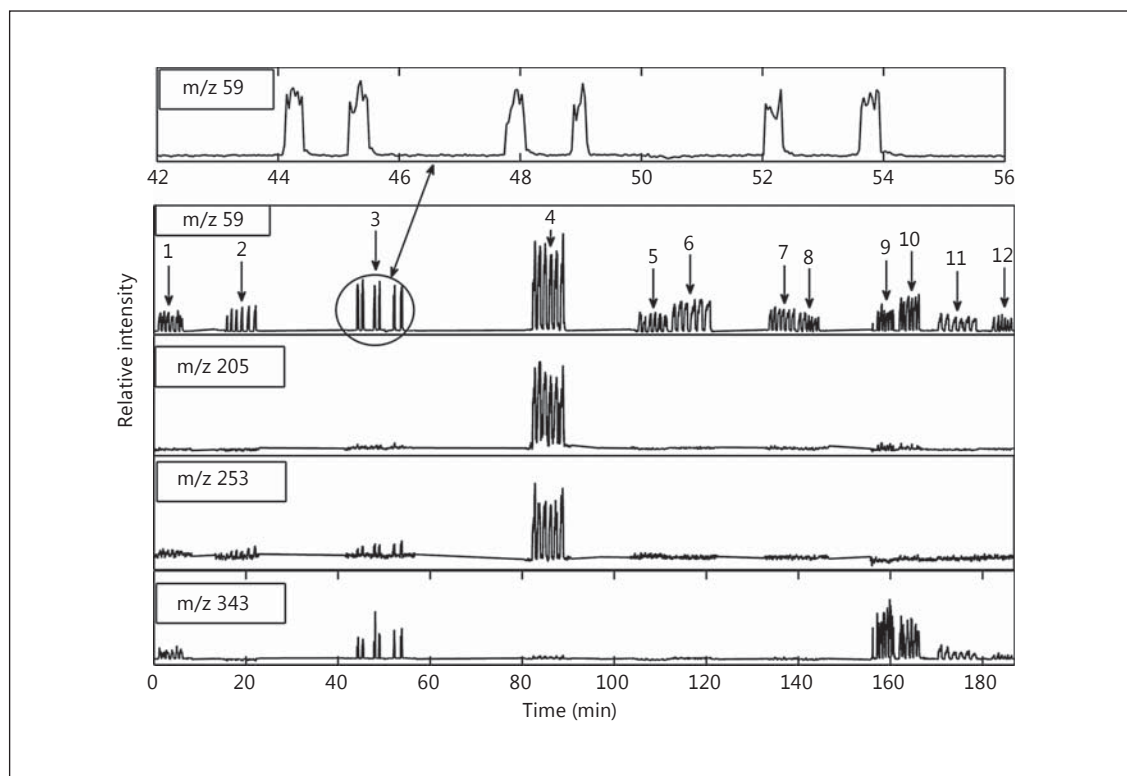
## Results

### Subjects

The baseline characteristics of the 4 subgroups are shown in tables 1 and 2. As expected, FEV<sub>1</sub> and FVC values in COPD patients were lower than in both smoking and nonsmoking controls. Patients with COPD were older, had smoked more pack-years than the controls and were being treated with various medications, including inhalational COPD therapy. However, the 2 COPD subgroups showed comparable characteristics.

### Mass-Spectral Breathprints

Figure 1 shows the typical ion intensity as a function of time for 4 selected  $m/z$  values present in the breath of a random selection of 12 subjects. Each trace ( $m/z$  59, 205, 253 and 343) corresponds to one compound present in breath for at least 1 of the 12 participants (labeled 1–12 at the top of the  $m/z$  59 trace). Subjects 2, 3 and 4 are COPD patients. Each subject breathed 6 consecutive times, the whole process taking typically  $< 10$  min/subject. For example, for the trace at  $m/z$  59, it is clearly observable how the ion intensity rises above the background level 6 times



**Fig. 1.** Time trace of the ion current for 4 selected  $m/z$  values (59, 201, 253 and 351) monitored in the breath of 12 randomly selected subjects. Individuals labeled as 2, 3 and 4 correspond to COPD patients (GOLD stages IV, IV and II, respectively), while the rest of the subjects shown in this example were nonsmoking controls.

Each subject provided 6 consecutive breath samples. The top trace shows a zoom at  $m/z$  59 (acetone) of the 6 replicate measurements of one of the COPD patients. It illustrates that excellent repeatability can be obtained even in cases of impaired breathing capacity. Note also how each subject shows a distinct breathprint.

for all 12 subjects. This particular compound corresponds to acetone, as it has been previously characterized [22]. For each individual, the height (intensity) of the acetone peaks is very similar for the 6 breaths, indicating a satisfactory repeatability (average of 12% RSD for the 61 participants) and robustness of the measurements, even for COPD patients (fig.1 – see zoom of acetone-subject 3), but clearly different among individuals [41].

#### Discrimination of Study Groups

The collected breath mass spectra were processed statistically with the aim of ultimately discriminate the study groups. Figure 2 (left) shows 3 box plots corresponding to 3 representative features found in this study to yield a  $p$  value of  $<0.05$  after a Mann-Whitney U test. The right plot shows a zoom of the mass spectrum in the region of interest with all the traces corresponding to each participant overlaid. The top peak, which corresponds to acetone, was found significantly increased in COPD versus

healthy smokers ( $p < 0.01$ ). The middle graph ( $m/z$  118) was found to be significantly ( $p < 0.005$ ) decreased in COPD patients' breath compared to the healthy non-smoking controls. The bottom graph shows an example of one peak found to be increased in GOLD stages III/IV vs. I/II ( $p < 0.05$ ). Tables 3–5 list the remaining peaks found in this study to be differently exhaled among the different groups under study. Figure 3 shows the corresponding principal component analysis plots for the signals listed in tables 3–5.

Finally, we tested whether these sets of exhaled compounds may be used to predict the presence/absence of disease and disease stage. Table 6 summarizes the classification results for the 3 binary comparisons performed, where 24/25 COPD patients and 8/11 smokers, 22/25 COPD patients and 23/25 nonsmokers and 12/13 GOLD stages I/II and 10/12 GOLD stages III/IV were correctly classified for each of the 3 leave-one-out cross-validations.

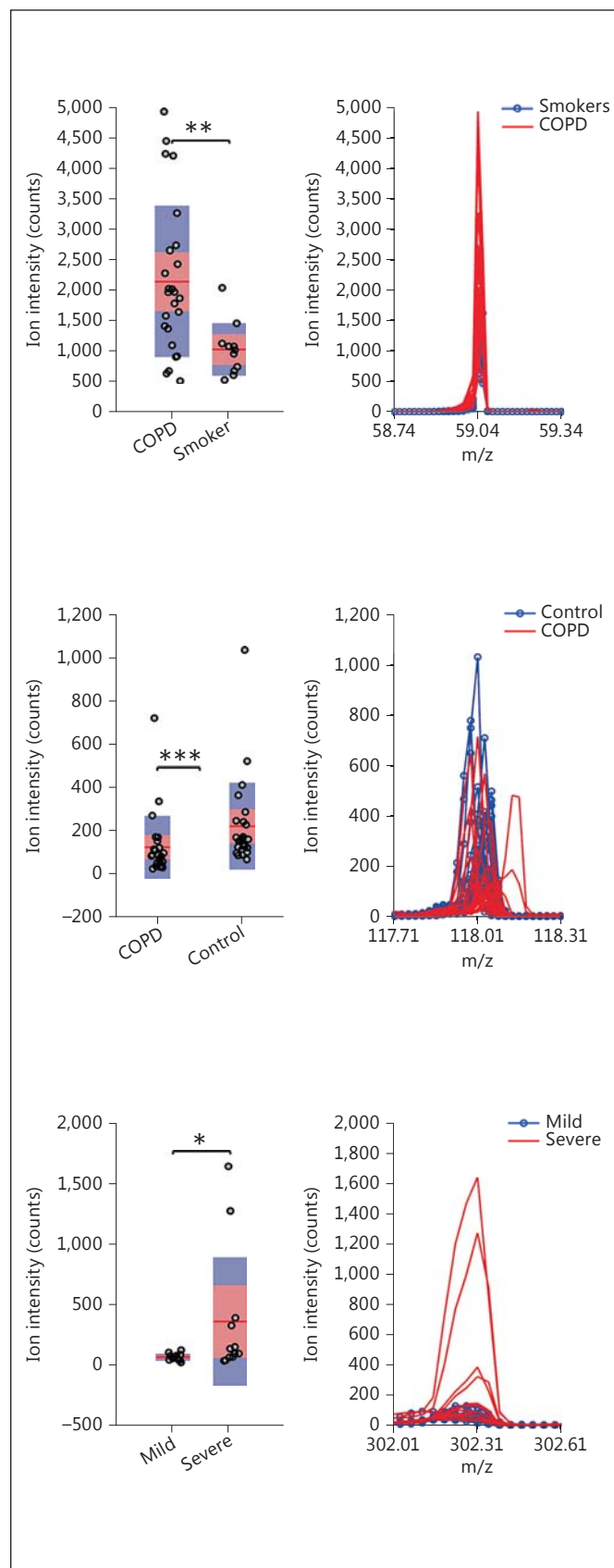


**Table 3.** Mass-spectral features found to be exhaled at significantly ( $p < 0.05$ ) different concentrations by COPD subjects and non-smoking controls

m/z	COPD subjects	Nonsmoking controls
79	63 (49–77)	102 (65–138)
102.1	134 (111–156)	191 (160–221)
118.0	120 (60–180)	217 (134–300)
143.1	931 (684–1,179)	1,218 (1,025–1,411)
157.1	664 (531–798)	844 (744–945)
158.1	557 (459–656)	723 (632–814)
160.1	214 (164–263)	335 (262–409)
174.1	338 (264–413)	391 (344–438)
180.1	541 (180–901)	623 (300–946)
185.1	493 (426–560)	580 (526–635)
185.9	62 (53–72)	77 (70–84)
186.2	496 (381–610)	956 (654–1,258)
188.1	284 (232–337)	357 (309–406)
190.1	301 (242–360)	393 (335–451)
202.2	268 (229–308)	416 (302–529)
227.1	722 (548–896)	509 (379–639)
240.2	152 (130–174)	216 (173–259)
245.2	228 (200–256)	193 (152–233)
247.2	386 (313–458)	320 (222–418)
258.1	161 (138–184)	203 (174–232)
265.1	129 (86–171)	114 (111–118)
268.2	95 (80–110)	119 (102–136)
271	110 (88–131)	78 (71–86)
271.3	912 (682–1,141)	614 (433–796)
277.1	160 (108–212)	107 (88–125)
287.3	218 (147–288)	125 (102–148)
327	419 (364–473)	484 (439–528)
328	54 (50–59)	62 (57–66)
329	416 (361–471)	477 (433–520)
329.3	111 (80–141)	77 (48–106)
330	60 (55–64)	67 (63–71)
331	148 (131–164)	165 (152–179)
332.2	71 (63–79)	95 (84–106)
344.1	270 (218–321)	362 (316–408)
345	58 (54–63)	67 (61–73)
346	256 (205–307)	359 (312–406)
348.1	103 (84–121)	134 (118–149)

The signal intensity values listed are means (95% confidence interval) in units of counts.

**Fig. 2.** Examples of compounds found to be exhaled at different concentrations between COPD subjects and smoking controls (top), COPD subjects and nonsmoking controls (middle) and those with COPD GOLD stages I/II and III/IV (bottom). The feature at m/z 59 corresponds to acetone and the one at m/z 118 to indole (most probable identity). \*  $p < 0.05$ , \*\*  $p < 0.01$  and \*\*\*  $p < 0.005$ .



**Table 4.** Mass-spectral features found to be exhaled at significantly ( $p < 0.05$ ) different concentrations by COPD subjects and smoking controls

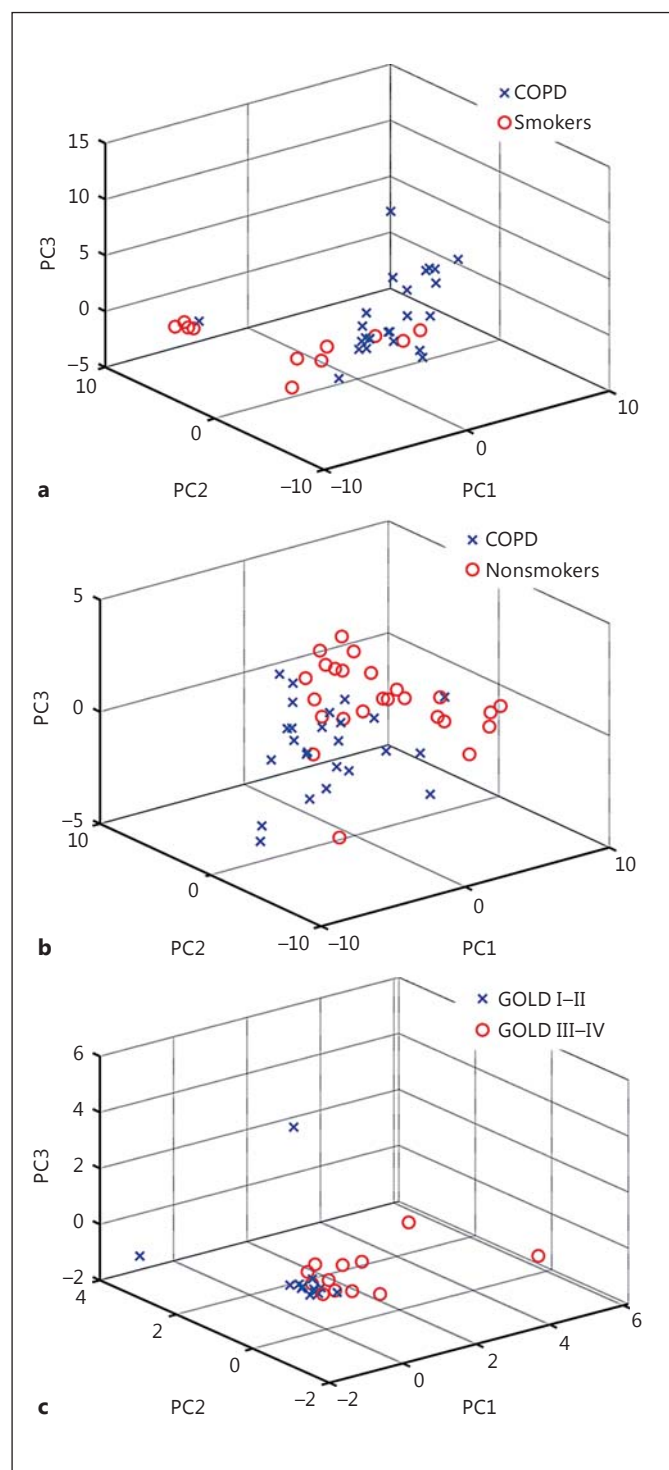
m/z	COPD subjects	Smoking controls	m/z	COPD subjects	Smoking controls
59	2,155 (1,644–2,666)	1,042 (752–1,332)	235.2	1,173 (949–1,397)	747 (173–1,321)
60.1	168 (136–200)	95 (78–111)	236.2	287 (218–355)	209 (70–348)
98	38 (27–50)	74 (36–112)	239.2	168 (142–194)	131 (109–153)
122	85 (70–99)	133 (105–160)	241.2	141 (123–158)	112 (79–145)
123.1	220 (98–342)	746 (219–1,274)	244.2	578 (461–695)	621 (–96 to 1,338)
124.1	86 (59–113)	191 (110–272)	247.2	386 (313–458)	287 (128–446)
126	71 (40–102)	223 (81–365)	249.2	279 (196–362)	189 (161–216)
146	132 (78–187)	335 (128–543)	251.2	186 (167–206)	150 (129–172)
150	111 (81–141)	241 (112–370)	253.1	138 (99–178)	106 (90–122)
155.1	1,280 (617–1,943)	468 (260–677)	255.1	258 (–17 to 532)	98 (89–107)
160.1	235 (164–306)	433 (252–614)	256.2	164 (78–250)	96 (69–122)
172.1	553 (411–695)	288 (102–475)	257.2	253 (–11 to 518)	84 (42–126)
182	68 (41–95)	186 (62–309)	258.1	161 (138–184)	121 (85–157)
189.1	195 (154–237)	132 (102–162)	267.2	172 (152–192)	129 (95–163)
191.1	366 (323–408)	271 (210–333)	269.2	104 (92–117)	76 (61–92)
194.1	105 (75–136)	181 (103–259)	271.3	912 (682–1,141)	357 (30–684)
195.2	1,856 (1,448–2,265)	1,065 (475–1,655)	272.3	265 (212–318)	146 (64–227)
198.1	167 (144–191)	115 (91–139)	273.2	96 (81–111)	66 (48–84)
199.1	435 (374–497)	282 (173–391)	276.2	131 (102–160)	246 (–121 to 612)
201.1	345 (284–407)	228 (179–278)	279.2	1,536 (1,310–1,762)	1,111 (839–1,384)
203.1	417 (326–509)	235 (118–352)	285.2	338 (53–623)	100 (57–143)
204.1	361 (302–421)	241 (132–351)	288.3	474 (363–586)	309 (69–550)
205.2	553 (158–948)	237 (137–338)	288.9	34 (28–39)	68 (43–93)
207.1	1,159 (924–1,394)	981 (19–1,943)	289.3	150 (121–179)	101 (43–159)
208.2	508 (394–621)	390 (35–745)	298.9	50 (40–60)	74 (51–97)
211.2	342 (288–395)	229 (166–293)	299.3	258 (136–381)	74 (48–101)
213.1	299 (244–354)	214 (166–263)	316.3	194 (100–289)	68 (32–103)
214	156 (132–181)	106 (70–142)	317.3	82 (59–105)	49 (37–62)
214.1	232 (195–270)	165 (123–208)	319.1	43 (34–53)	55 (42–68)
215.1	272 (236–308)	166 (111–221)	323.1	37 (30–44)	48 (36–59)
218.1	294 (243–345)	216 (135–298)	325.1	36 (28–45)	60 (38–82)
222.2	224 (181–266)	275 (–49 to 599)	329.3	111 (80–141)	45 (10–80)
223.1	1,390 (1,109–1,670)	954 (833–1,075)	330.3	65 (52–79)	34 (18–50)
224.1	317 (250–384)	242 (210–273)	The signal intensity values are listed as means (95% confidence interval) in units of counts.		
227.1	722 (548–896)	538 (–29 to 1,105)			
228.2	215 (181–250)	186 (36–335)			

**Table 5.** Mass spectral features found to be exhaled at significantly ( $p < 0.05$ ) different concentrations by both COPD subgroups

m/z	GOLD I/II	GOLD III/IV
165	277 (210–344)	210 (178–242)
286.3	70 (56–84)	228 (6–451)
302.3	63 (46–80)	358 (21–695)
308.2	53 (17–90)	55 (41–70)
The signal intensity values listed are means (95% confidence interval) in units of counts.		

**Table 6.** Summary of the leave-one-out cross-validation for the 3 classifications

	COPD (n = 25) vs. smokers (n = 11)	COPD (n = 25) vs. nonsmokers (n = 25)	GOLD I/II (n = 13) vs. GOLD III/IV (n = 12)
Sensitivity	96.0%	88.0%	92.3%
Specificity	72.7%	92.0%	83.3%
Positive predictive value	88.9%	91.7%	85.7%
Negative predictive value	88.9%	88.5%	90.9%



**Fig. 3.** 3-Dimensional principal component analysis (PCA) plot of the mass-spectral breathprints containing the most significant features of COPD subjects and smoking controls (a), COPD subjects and nonsmoking controls (b) and subjects with COPD GOLD stages I/II and III/IV (c). As expected, clustering according to health status is more obvious in the COPD vs. healthy controls (a, b) than in the COPD subgroups (c).

## Discussion

The analysis of exhaled breath is believed to hold promise in the study of lung diseases since exhaled air is in direct contact with the respiratory tract and therefore may carry relevant biological markers. For this reason, quite some efforts have been devoted over the last years to develop instrumentation and methods capable of detecting exhaled biomarkers for a number of lung diseases, including COPD [42].

For example, the analysis of exhaled-breath condensate, covering molecules ranging from small metabolites to large proteins [3], has been shown to shed light on the disease phenotype [5, 7]. The main drawback of analyzing exhaled-breath condensate is likely associated with the difficulty to standardize sample collection. For this reason, new sample collection techniques are currently under development and investigation [13, 43].

Another technology targeting exhaled breath is the so-called electronic nose, which responds to combinations of volatiles present in breath providing breathprints amenable for statistical analysis [15, 16]. However, this technology is not suitable to provide structural elucidation of the breath compounds providing separation among different cohorts, which is crucial for gaining insight into underlying biochemical COPD mechanisms. On the other hand, an advantageous feature is the portability of these devices. Similarly, gas chromatography ion mobility spectrometry provides good portability and pattern identification, but lacks unambiguous structural identification capabilities [14, 17].

Complementarily to these techniques, breath can be analyzed by MS, either off-line (e.g. gas chromatography MS [15, 44]) or in real time by a number of different techniques. The main advantage of MS is its selectivity (especially high-resolution instruments), which enables a wide analyte coverage as well as the structural identification of the components detectable in breath. There are mainly three different approaches to analyze breath in real time: proton transfer reaction MS [45], selected ion flow tube MS [46] and SESI-MS.

In real-time techniques (i.e. no sample collection-manipulation), the risks of introducing artifacts due to sample handling are minimized. Not only is the acquisition of the breathprint data fast, but also the complete processing and classification of the spectra afterwards can be performed rapidly. For example, the whole process including the 6 replicate breath analyses and the statistical classification (diagnosis) can be accomplished in 10 min. Currently, the main drawback of SESI-MS, as



opposed to electronic noses, for example, is the lack of portability.

All the real-time MS techniques available nowadays can detect compounds of relatively high volatility, e.g. acetone, which is a well-known metabolite present in breath [47, 48] (fig. 1, 2,  $m/z$  59). However, the excellent sensitivity of SESI-MS (down to parts-per-trillion [24, 49]) offers the opportunity to target a set of molecules never observed before in breath in real time [22–24, 50]. A recent model has proposed that ionization at atmospheric pressure leads to higher ionization efficiencies, which may explain SESI-MS efficiency in this relatively high  $m/z$  range [51]. Notice, for example, that species with molecular weights as large as  $m/z$  343 are clearly detectable (fig. 1). Another example is the data in tables 3–5, where masses well beyond  $m/z$  200 (the typical  $m/z$  limit of other online techniques) strongly contribute to discerning COPD patients and controls. Indeed, it is reasonable to hypothesize that more complex (heavier) biomolecules may carry more accurate information since they may be linked to just one particular biochemical process, rather than to several.

In contrast with electronic noses, for example, one major advantage of MS is that it allows for the chemical characterization of the detected exhaled molecules. For example, by further fragmentation measurements, the ion seen at  $m/z$  59 was assigned to acetone [22], which is an intermediate of the synthesis and degradation of ketone bodies and propanoate metabolism ([http://www.kegg.jp/dbget-bin/www\\_bget?cpd:C00207](http://www.kegg.jp/dbget-bin/www_bget?cpd:C00207)). In this study, acetone was found to be exhaled in lower concentrations in the smokers group than among the COPD patients (fig. 2).

In the second phase of our study, we had access to a high-resolution instrument with fragmentation capabilities (Waters' Synapt G2S). While the identification with high confidence of the large number of compounds found in breath via SESI-MS is ongoing, we anticipate here that, based on accurate mass, isotopic pattern and MS/MS data, the peak at  $m/z$  118 corresponds to indole. It was also tentatively identified in breath via selected ion flow tube MS [52] and it has been unambiguously detected offline (gas chromatography MS). Recently, it was found to be increased in humans who are under stress [53], and in patients suffering liver malfunction compared to a healthy cohort, it was found to be decreased [54]. Indole is one metabolite involved in tryptophan metabolism ([http://www.kegg.jp/dbget-bin/www\\_bget?cpd:C00463](http://www.kegg.jp/dbget-bin/www_bget?cpd:C00463)) and binds to albumin in blood [55]. At present, the reasons why indole may be decreased in COPD patients' breath

are unclear and this finding is under further investigation for confirmation.

Another interesting observation made is that the combination of the large number of different components detectable in breath by SESI-MS provides a sort of unique 'fingerprint' for each individual, observable even by visual inspection (fig. 1). The existence of individual metabolic phenotypes, which seem to be stable during prolonged periods, has been hypothesized and experimental evidence has been published [56]. Previous [24] and recent [41] investigations by our group support this hypothesis via the analysis of exhaled breath. This finding has important implications in personalized healthcare [57] since, for example, it is conceivable to monitor health status through breath in an individualized fashion (e.g. monitoring of COPD exacerbation episodes and response to medication).

Our study has some limitations. At this exploratory stage, we selected a heterogeneous control group involving young healthy smokers and nonsmokers, with the aim that differences among the groups be readily observable. This was indeed the case, and therefore current efforts are being devoted to confirm if the satisfactory classifications obtained here may be achievable in age-matched control groups. Note that both COPD subgroups studied above had a comparable mean age and that even so, they could be classified (table 6). This suggests that the observed differences between groups indeed stem from states of disease/nondisease rather than from age differences. Similarly, the COPD patients were on medications for COPD and comorbidities. As shown previously [58], our technique seems suitable for monitoring pharmacokinetics through exhaled breath. Thus, it is conceivable that medications may be potential confounders [59]. However, except for a higher number of patients using inhaled steroids in the COPD GOLD III/IV subgroup than in the GOLD I/II subgroup, these subgroups were on similar medications and we were able to detect differences in their breathprints (fig. 2, 3; table 5). Despite this limitation, and given the precedent on antiepileptic drug monitoring in breath [58], we hypothesize that this technique might also be useful for a better understanding of the mechanism of action of drugs for COPD [60–62].

To our knowledge, this is the first real-time MS study investigating the exhaled breath of COPD patients. We conclude that exhaled breath can be analyzed in real time covering masses well beyond 300 Da. We collected 6 replicate measurements per subject in <10 min, on average, with excellent repeatability and with a strong individual-dependent response. A subset of mass-spectral features

discriminating COPD (all stages; n = 25) versus healthy nonsmokers (n = 25), COPD (all stages; n = 25) versus healthy smokers (n = 11) and mild COPD (GOLD I/II; n = 13) versus severe COPD (GOLD III/IV; n = 12) were identified. These exhaled molecules were satisfactory predictors of the presence/absence of disease and an indication of disease stage.

This study provides a qualitative indication, suggesting that real-time breath analysis could add valuable information to the current state-of-the-art in COPD diagnosis and phenotyping. While it is premature to conclude

that this technique will be used routinely in the future to diagnose COPD as well as to understand its underlying mechanisms, these exploratory measurements are encouraging for the further pursuit of this approach.

## Acknowledgements

This research was supported by a Marie Curie European Reintegration Grant (PMLS) within the 7th European Community Framework Programme (276860). René Dreier is gratefully acknowledged for constructing the sampling device.

## References

- 1 Saetta M, Turato G, Maestrelli P, Mapp C, Fabbri L: Cellular and structural bases of chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2001;163:1304–1309.
- 2 Cazzola M, Novelli G: Biomarkers in COPD. *Pulm Pharmacol Ther* 2010;23:493–500.
- 3 O'Reilly P, Bailey W: Clinical use of exhaled biomarkers in COPD. *Int J Chron Obstr Pulm Dis* 2007;2:403–408.
- 4 Kharitonov SA, Barnes PJ: Exhaled biomarkers. *Chest* 2006;130:1541–1546.
- 5 Montuschi P: Exhaled breath condensate analysis in patients with COPD. *Clin Chim Acta* 2005;356:22–34.
- 6 Biernacki WA, Kharitonov SA, Barnes PJ: Increased leukotriene b4 and 8-isoprostane in exhaled breath condensate of patients with exacerbations of COPD. *Thorax* 2003;58:294–298.
- 7 de Laurentiis G, Paris D, Melck D, Maniscalco M, Marsico S, Corso G, Motta A, Sofia M: Metabonomic analysis of exhaled breath condensate in adults by nuclear magnetic resonance spectroscopy. *Eur Respir J* 2008;32:1175–1183.
- 8 Montuschi P, Paris D, Melck D, Lucidi V, Ciabattini G, Raia V, Calabrese C, Bush A, Barnes PJ, Motta A: NMR spectroscopy metabolomic profiling of exhaled breath condensate in patients with stable and unstable cystic fibrosis. *Thorax* 2012;67:222–228.
- 9 Motta A, Paris D, Melck D, de Laurentiis G, Maniscalco M, Sofia M, Montuschi P: Nuclear magnetic resonance-based metabolomics of exhaled breath condensate: methodological aspects. *Eur Respir J* 2012;39:498–500.
- 10 Lucidi V, Ciabattini G, Bella S, Barnes PJ, Montuschi P: Exhaled 8-isoprostane and prostaglandin E(2) in patients with stable and unstable cystic fibrosis. *Free Radic Biol Med* 2008;45:913–919.
- 11 Montuschi P, Barnes PJ, Ciabattini G: Measurement of 8-isoprostane in exhaled breath condensate. *Methods Mol Biol* 2010;594:73–84.
- 12 Malerba M, Montuschi P: Non-invasive biomarkers of lung inflammation in smoking subjects. *Curr Med Chem* 2012;19:187–196.
- 13 Almstrand AC, Josefson M, Bredberg A, Lausmaa J, Sjövall P, Larsson P, Olin AC: TOF-SIMS analysis of exhaled particles from patients with asthma and healthy controls. *Eur Respir J* 2012;39:59–66.
- 14 Basanta M, Jarvis RM, Xu Y, Blackburn G, Tal-Singer R, Woodcock A, Singh D, Goodacre R, Paul Thomas CL, Fowler SJ: Non-invasive metabolomic analysis of breath using differential mobility spectrometry in patients with chronic obstructive pulmonary disease and healthy smokers. *Analyst* 2010;135:315–320.
- 15 Fens N, De Nijs SB, Peters S, Dekker T, Knobel HH, Vink TJ, Willard NP, Zwiderman AH, Krouwelse FH, Janssen HG, Lutter R, Sterk PJ: Exhaled air molecular profiling in relation to inflammatory subtype and activity in COPD. *Eur Respir J* 2011;38:1301–1309.
- 16 Fens N, Zwiderman AH, van der Schee MP, de Nijs SB, Dijkers E, Roldaan AC, Cheung D, Bel EH, Sterk PJ: Exhaled breath profiling enables discrimination of chronic obstructive pulmonary disease and asthma. *Am J Respir Crit Care Med* 2009;180:1076–1082.
- 17 Westhoff M, Litterst P, Maddula S, Bödeker B, Rahmann S, Davies A, Baumbach J: Differentiation of chronic obstructive pulmonary disease (COPD) including lung cancer from healthy control group by breath analysis using ion mobility spectrometry. *Int J Ion Mobil Spectrom* 2010;13:131–139.
- 18 Van Berkel JBN, Dallinga JW, Möller GM, Godschalk RWL, Moonen EJ, Wouters EFM, Van Schooten FJ: A profile of volatile organic compounds in breath discriminates COPD patients from controls. *Respir Med* 2010;104:557–563.
- 19 Montuschi P, Mores N, Trove A, Mondino C, Barnes PJ: The electronic nose in respiratory medicine. *Respiration* 2013;85:72–84.
- 20 Bofan M, Mores N, Baron M, Dabrowska M, Valente S, Schmid M, Trove A, Conforto S, Zini G, Cattani P, Fuso L, Mautone A, Mondino C, Pagliari G, D'Alessio T, Montuschi P: Within-day and between-day repeatability of measurements with an electronic nose in patients with COPD. *J Breath Res* 2013;7:017103.
- 21 Ehmann R, Boedeker E, Friedrich U, Sagert J, Dippon J, Friedel G, Walles T: Canine scent detection in the diagnosis of lung cancer: revisiting a puzzling phenomenon. *Eur Respir J* 2012;39:669–676.
- 22 Martínez-Lozano P, de la Mora JF: Electrospray ionization of volatiles in breath. *Int J Mass Spectrom* 2007;265:68–72.
- 23 Martínez-Lozano P, Fernández de la Mora J: Direct analysis of fatty acid vapors in breath by electrospray ionization and atmospheric pressure ionization-mass spectrometry. *Anal Chem* 2008;80:8210–8215.
- 24 Martínez-Lozano P, Zingaro L, Finiguerra A, Cristoni S: Secondary electrospray ionization-mass spectrometry: breath study on a control group. *J Breath Res* 2011;5:016002.
- 25 Martínez-Lozano Sinues P, Kohler M, Zenobi R: Monitoring diurnal changes in exhaled human breath. *Anal Chem* 2013;85:369–373.
- 26 Martínez-Lozano P, Fernández de la Mora J: Direct analysis of fatty acid vapors in breath by electrospray ionization and atmospheric pressure ionization-mass spectrometry. *Analytical Chemistry* 2008;80:8210–8215.
- 27 Martínez-Lozano P, Fernández de la Mora J: Electrospray ionization of volatiles in breath. *Int J Mass Spectrom* 2007;265:68–72.
- 28 Berchtold C, Meier L, Steinhoff R, Zenobi R: A new strategy based on real-time secondary electrospray ionization and high-resolution mass spectrometry to discriminate endogenous and exogenous compounds in exhaled breath. *Metabolomics* 2013, DOI: 10.1007/s11306-013-0568-z.
- 29 Wu C, Siems WF, Hill HH: Secondary electrospray ionization ion mobility spectrometry/mass spectrometry of illicit drugs. *Anal Chem* 2000;72:396–403.
- 30 Zhu J, Bean HD, Jiménez-Díaz J, Hill JE: Secondary electrospray ionization-mass spectrometry (SESI-MS) breathprinting of multiple bacterial lung pathogens, a mouse model study. *J Appl Physiol* 2013;114:1544–1549.

- 31 Jiangjiang Z, Heather DB, Matthew JW, Laurie WL, Jane EH: Detecting bacterial lung infections: in vivo evaluation of in vitro volatile fingerprints. *J Breath Res* 2013;7:016003.
- 32 Jiangjiang Z, Jaime J-D, Heather DB, Nirav AD, Minara IA, Lennart KAL, Jane EH: Robust detection of *P. aeruginosa* and *S. aureus* acute lung infections by secondary electrospray ionization-mass spectrometry (SESI-MS) breathprinting: from initial infection to clearance. *J Breath Res* 2013;7:037106.
- 33 Zhu J, Bean HD, Wargo M, Leclair L, Hill JE: Detecting bacterial lung infections: in vivo evaluation of in vitro volatile fingerprints. *J Breath Res* 2013;7:016003.
- 34 Bean HD, Zhu J, Hill JE: Characterizing bacterial volatiles using secondary electrospray ionization mass spectrometry (SESI-MS). *J Vis Exp* 2011:e2664.
- 35 Westhoff M, Litterst P, Freitag L, Urfer W, Bader S, Baumbach JI: Ion mobility spectrometry for the detection of volatile organic compounds in exhaled breath of patients with lung cancer: results of a pilot study. *Thorax* 2009;64:744–748.
- 36 National Institutes of Health: Global Initiative for Chronic Obstructive Lung Disease. NHLBI/WHO Workshop Report 2001 (pub No. 2701).
- 37 Martinez-Lozano Sinues P, Criado E, Vidal G: Mechanistic study on the ionization of trace gases by an electrospray plume. *Int J Mass Spectrom* 2012;313:21–29.
- 38 Vidakovic B: Statistics for Bioengineering Sciences: With MATLAB and WinBUGS Support. New York, Springer, 2011.
- 39 van den Berg R, Hoefsloot H, Westerhuis J, Smilde A, van der Werf M: Centering, scaling, and transformations: improving the biological information content of metabolomics data. *BMC Genom* 2006;7:142.
- 40 Jackson JE: A User's Guide to Principal Components. Hoboken, Wiley, 2005.
- 41 Martinez-Lozano Sinues P, Kohler M, Zenobi R: Human breath analysis may support the existence of individual metabolic phenotypes. *PLoS ONE* 2013;8:e59909.
- 42 Barnes PJ: Future perspectives on exhaled biomarkers. *Eur Respir Soc J* 2010;49:237–246.
- 43 McKenzie JH, McDevitt JJ, Fabian MP, Hwang GM, Milton DK: Collection of aerosolized human cytokines using Teflon® filters. *PLoS ONE* 2012;7:e35814.
- 44 Phillips CO, Syed Y, Parthalain NM, Zwigelaar R, Claypole TC, Lewis KE: Machine learning methods on exhaled volatile organic compounds for distinguishing COPD patients from healthy controls. *J Breath Res* 2012;6:036003.
- 45 Lindinger W, Hansel A, Jordan A: On-line monitoring of volatile organic compounds at PPTV levels by means of proton-transfer-reaction mass spectrometry (PTR-MS) medical applications, food control and environmental research. *Int J Mass Spectrom Ion Processes* 1998;173:191–241.
- 46 Smith D, Španěl P: Selected ion flow tube mass spectrometry (SIFT-MS) for on-line trace gas analysis. *Mass Spectrom Rev* 2005;24:661–700.
- 47 Spanel P, Dryahina K, Rejskova A, Chippendale TWE, Smith D: Breath acetone concentration; biological variability and the influence of diet. *Physiol Meas* 2011;32:N23–N31.
- 48 Schwarz K, Pizzini A, Arendacká B, Zerlauth K, Filipiak W, Schmid A, Dzien A, Neuner S, Lechleitner M, Scholl-Bürgi S, Miekisch W, Schubert J, Unterkofler K, Witkovský V, Gastl G, Amann A: Breath acetone – aspects of normal physiology related to age and gender as determined in a PTR-MS study. *J Breath Res* 2009;3:027003.
- 49 Martínez-Lozano P, Rus J, Fernández de la Mora G, Hernández M, Fernández de la Mora J: Secondary electrospray ionization (SESI) of ambient vapors for explosive detection at concentrations below parts per trillion. *J Am Soc Mass Spectrom* 2009;20:287–294.
- 50 Chen H, Wortmann A, Zhang W, Zenobi R: Rapid in vivo fingerprinting of nonvolatile compounds in breath by extractive electrospray ionization quadrupole time-of-flight mass spectrometry. *Angew Chem Int Ed Engl* 2007;46:580–583.
- 51 Fernandez de la Mora J: Ionization of vapor molecules by an electrospray cloud. *Int J Mass Spectrom* 2011;300:182–193.
- 52 Ross B: Sub-parts per billion detection of trace volatile chemicals in human breath using selected ion flow tube mass spectrometry. *BMC Res Notes* 2008;1:41.
- 53 Turner MA, Bandelow S, Edwards L, Patel P, Martin HJ, Wilson ID, Thomas CLP: The effect of a paced auditory serial addition test (PASAT) intervention on the profile of volatile organic compounds in human breath: a pilot study. *J Breath Res* 2013;7:017102.
- 54 Van den Velde S, Nevens F, Van hee P, van Steenberghe D, Quirynen M: GC-MS analysis of breath odor compounds in liver patients. *J Chromatogr B* 2008;875:344–348.
- 55 Gambhir KK, McMenamy RH, Watson F: Positions in human serum albumin which involve the indole binding site. Sequence of 107-residue fragment. *J Biol Chem* 1975;250:6711–6719.
- 56 Assfalg M, Bertini I, Colangioli D, Luchinat C, Schäfer H, Schütz B, Spraul M: Evidence of different metabolic phenotypes in humans. *Proc Natl Acad Sci* 2008;105:1420–1424.
- 57 Montuschi P: Toward a personalized pharmacotherapy of respiratory diseases. *Front Pharmacol* 2010;1:131.
- 58 Gamez G, Zhu L, Disko A, Chen H, Azov V, Chingin K, Kramer G, Zenobi R: Real-time, in vivo monitoring and pharmacokinetics of valproic acid via a novel biomarker in exhaled breath. *Chem Commun* 2011;47:4884–4886.
- 59 Montuschi P: Pharmacological treatment of chronic obstructive pulmonary disease. *Int J Chron Obstr Pulmon Dis* 2006;1:409.
- 60 Martinez-Lozano Sinues P, Zenobi R, Kohler M: Analysis of the exhalome: a diagnostic tool of the future. *Chest J* 2013;144:746–749.
- 61 Fuso L, Mores N, Valente S, Malerba M, Montuschi P: Long-acting beta-agonists and their association with inhaled corticosteroids in COPD. *Curr Med Chem* 2013;20:1477–1495.
- 62 Montuschi P, Macagno F, Valente S, Fuso L: Inhaled muscarinic acetylcholine receptor antagonists for treatment of COPD. *Curr Med Chem* 2013;20:1464–1476.